

# Introducing the Archivematica digital preservation system

Evelyn McLellan

Presentation for the Metropolitan New York  
Library Council, January 18, 2013



Artefactual Systems is located in Greater Vancouver on Canada's west coast. The easiest way to reach us is via email. We work with [clients](#) worldwide:



[View Larger Map](#)

[Street View](#)



We are one of the most hard-working and respected technical crews in the archives and library community. We are archivists, librarians and developers that are passionate about open-source and standards-based systems. We put out and support several [Archivematica](#) and [AtoM](#) releases a year and deliver [technical services](#) to a growing list of satisfied [clients](#).

**Peter Van Garderen**

President

**Evelyn McLellan**

Director, Consulting Services

**David Juhasz**

Director, Technical Services

**Courtney C. Mumma**

Archivematica Product Manager

**Jessica Bushey**

AtoM Product Manager

**Jesús García Crespo**

Software Developer

**Joseph Perry**

Software Developer

**Austin Trask**

Systems Technician

**Mike Cantelon**

Software Developer

**Dan Gillean**

Systems Analyst

**Justin Simpson**

Software Developer

**Mike Gale**

Software Developer

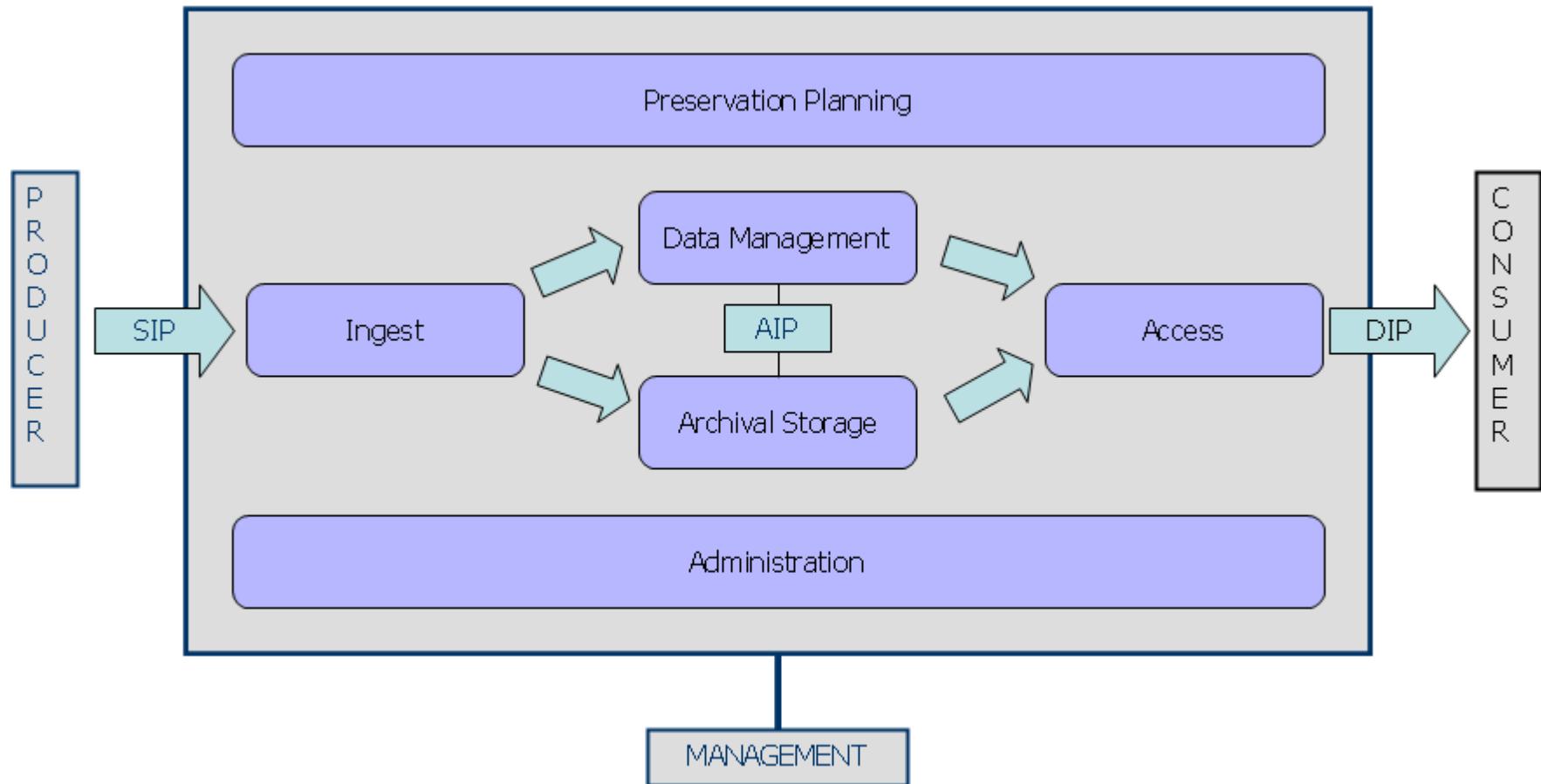
# Digital preservation: the problem

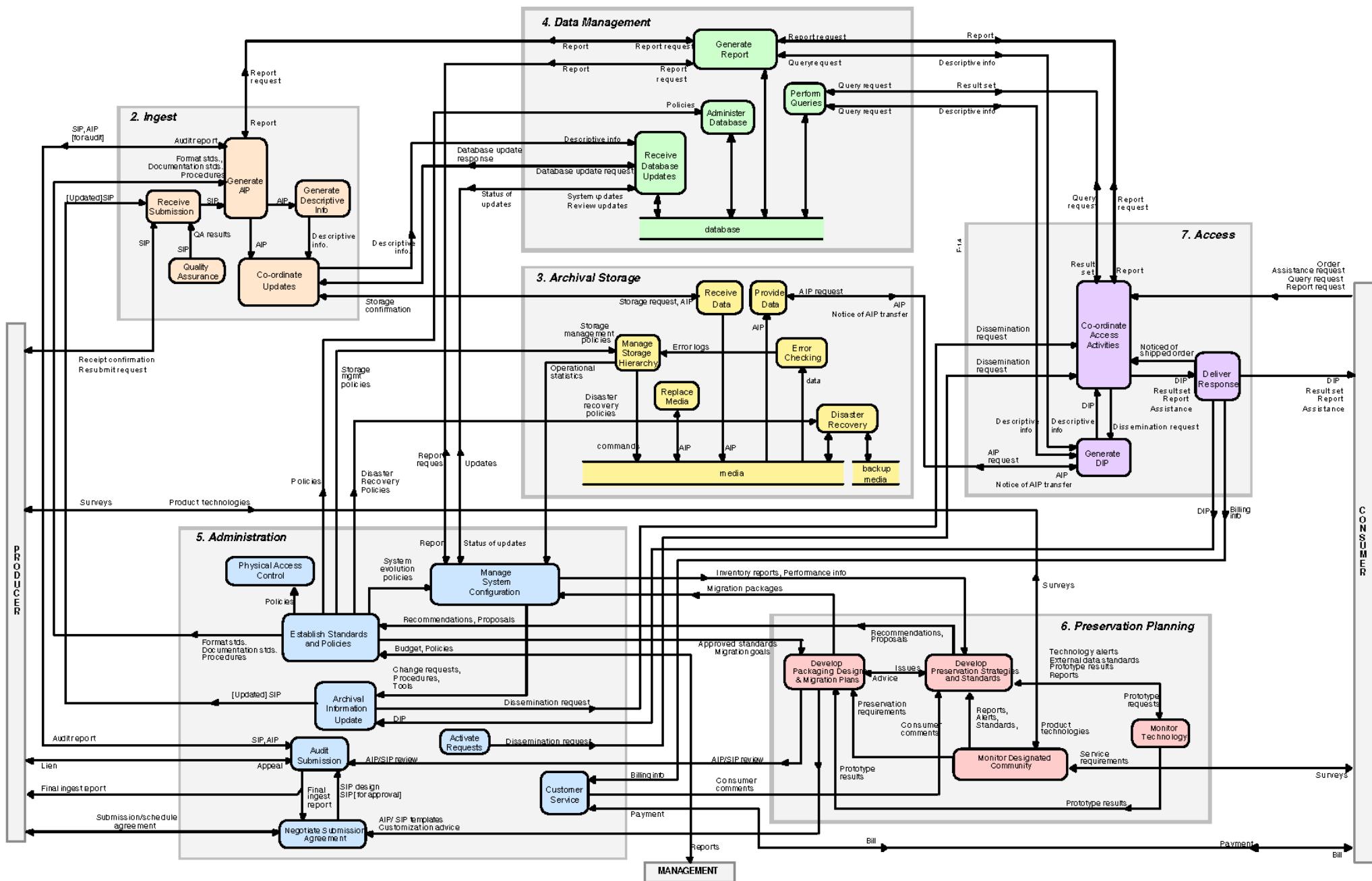
Digital objects are inherently fragile: they are easily lost or corrupted

Hardware, software and file formats become obsolete

Digital objects can lose their identity, context and meaning

# ISO 14721: Open Archival Information System





# OAIS definitions

- Submission Information Package (SIP)
  - A body of digital objects and associated metadata transferred from the producer
- Archival Information Package (AIP)
  - A package derived from the SIP containing the digital objects and associated metadata that is preserved in the digital preservation system
- Dissemination Information Package (DIP)
  - A package derived from the AIP containing digital objects and associated metadata delivered to the consumer

# What does OAIS tell us to do?

- Ingest a SIP
  - Accept a SIP from a Producer
  - Prepare an AIP
- Preserve the AIP
  - Ensure that the objects are securely stored
  - Ensure the ongoing ability to access and use the objects
- Manage information about the objects
- Disseminate the objects

# What is Archivematica?

Free and open-source digital preservation system

Designed to maintain standards-based, long-term  
access to collections of digital objects

# What is Archivematica?

Allows users to process digital objects from ingest to access consistent with the ISO-OAIS functional model

Creates a simple, highly interoperable AIP using recognized standards such as METS, PREMIS, Dublin Core, LOC BagIt format

Implements format normalization upon ingest and preserves originals to support emulation and migration strategies

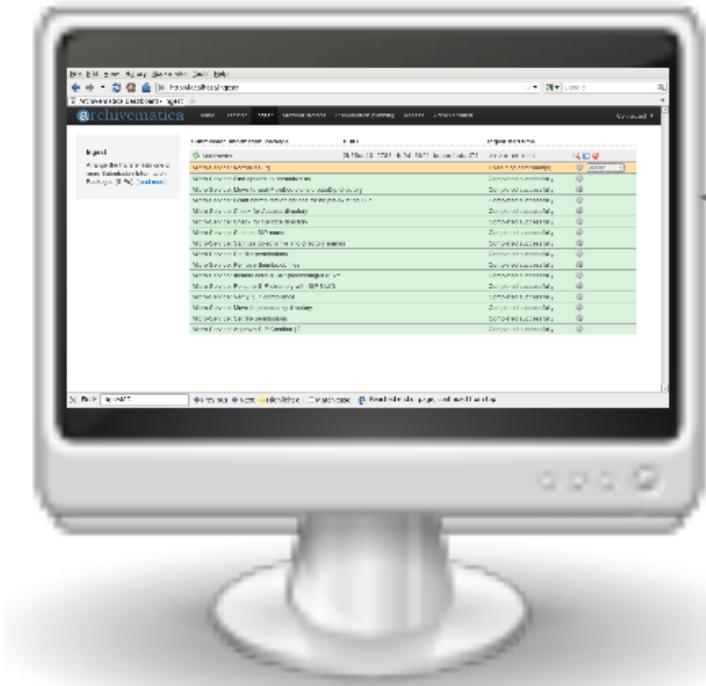
# What is Archivematica?

Consists of an integrated suite of open-source software tools

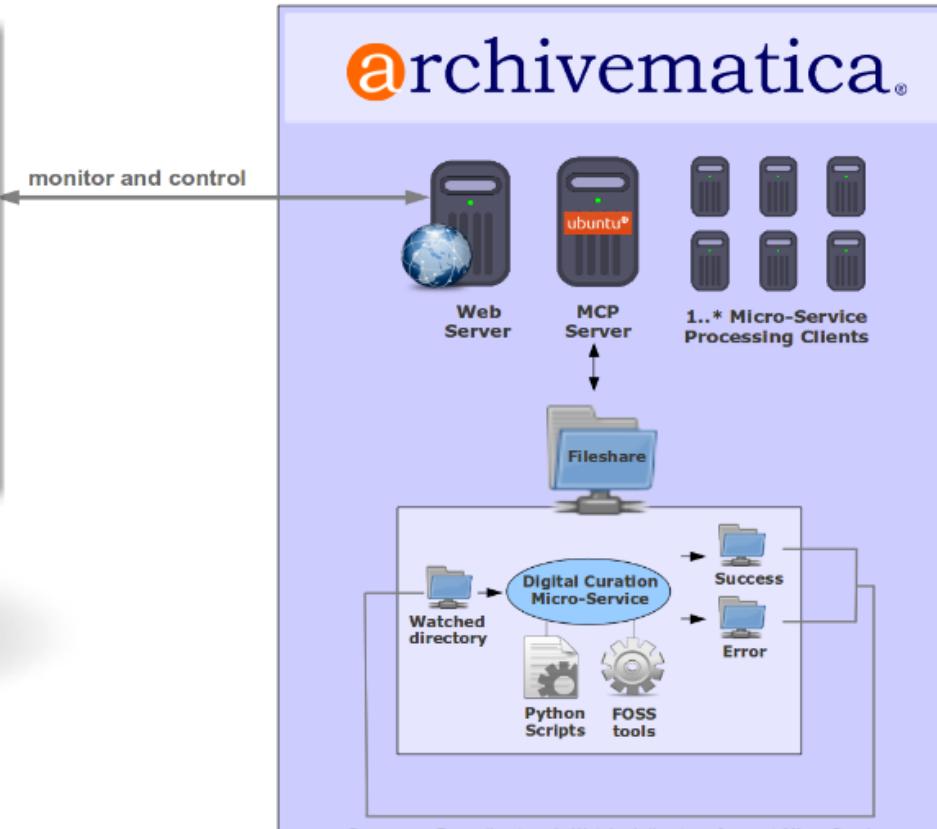
Processes digital objects through a pipeline of micro-services provided by the various bundled software tools

Allows for user interaction with the system through a web-based dashboard

Purpose is to create an AIP which can be stored and managed over time



MANAGEMENT  
Web Dashboard



or

transfer of  
digital files  
& metadata



# Micro-services

Ingestion

Fixity check

UUID assignment

Message digest calculation  
(checksum assignment)

Unpacking

Virus check

Format identification

Format validation

SIP preparation

Format normalization

AIP preparation

AIP packaging and  
compression

AIP storage

Generation of DIP

DIP upload

# The AIP METS file

<dmdSec>

Dublin Core

Ingested descriptive metadata

<amdSec>

PREMIS object metadata

PREMIS events

PREMIS rights

PREMIS agents

<fileSec>

List files in the AIP and their roles (original, preservation,  
submission documentation)

 archivematica.

Links files to their amdSecs

 artefactual  
systems inc.

# Ongoing development

Format policy register (FPR)

Bulk metadata import

Manual normalization workflow

Dataset preservation

Email preservation

Advanced AIP search

OAI harvesting/AIP uploading and versioning

Scalability testing

System integration: DSpace, Fedora, LOCKSS,  
Archivists Toolkit, CONTENTdm...

 archivematica.